IN THE CLAIMS

This listing of claims will replace all prior versions, and listings, of claims in the application. An identifier indicating the status of each claim is provided.

Listing of Claims

1. (Currently Amended)   A video/audio signal processing method for processing supplied compression-encoded video/audio signals, said method comprising the steps of:

parsing said video/audio signals in a compressed domain of the video/audio signals and extracting therefrom motion vectors of said video/audio signals, DCT-coefficients and macroblock-type;

using said extracted motion vectors, DCT-coefficients and macroblock-type to extract at least one compressed domain feature point representing characteristics of said video/audio signals in a compressed domain of said video/audio signals;

performing motion estimation of the extracted feature points; and

tracking the feature points associated with a motion vector through a pre-set number of frames of said video/audio signals,

wherein said extraction step includes a step of calculating the block relevance metric of all blocks according to said DCT-coefficients in the current frame to determine a block having high relevance as a candidate of the feature point selected as the next feature point based on said motion estimation step.

2. (Canceled)

3. (Previously Presented)   The video/audio processing method according to claim 1, wherein said extraction step includes a step of performing inverse transform of transforming said compressed domain only for the blocks of high relevance selected by said metric calculating step and of performing motion compensation for a prediction coded macroblock or a bidirectionally prediction coded macroblock.

4. (Original)   The video/audio processing method according to claim 3, wherein said inverse transform is inverse discrete cosine transform.

5. (Original)   The video/audio processing method according to claim 2, wherein said current frame includes an arbitrarily shaped video object plane.

6. (Original)   The video/audio processing method according to claim 4 including calculating and extracting a block signature for the current block of high relevance as selected in a discrete cosine transform domain using part or all of discrete cosine transform coefficients in a block.

7. (Original)   The video/audio processing method according to claim 4 including calculating and extracting a block signature for the current block of high relevance as selected in a discrete cosine transform domain using part or all of individually weighted discrete cosine transform coefficients in a block.

8. (Original) The video/audio processing method according to claim 4 including calculating a block signature for the current block of high relevance as selected in a pel domain.

9. (Original) The video/audio processing method according to claim 1, wherein said motion estimation step includes a step of calculating an estimated motion vector, the position of a reference block and a search area in a reference frame.

10. (Original) The video/audio processing method according to claim 9 including applying inverse transform of transforming said compressed domain to all blocks in an intra-macroblock in a search area of a reference frame.

11. (Original) The video/audio processing method according to claim 10, wherein said inverse transform is inverse discrete cosine transform.

12. (Original) The video/audio processing method according to claim 11 including performing inverse IDCT and motion compensation on all blocks in a prediction coded macroblock or in a bidirectional prediction coded macroblock in a search area of a reference frame.

13. (Original) The video/audio processing method according to claim 9, wherein said motion estimation step and said feature point tracking step include

a step of performing motion prediction or feature point tracking in a pel area for all search locations in the reference frame around the predicted motion vector in order to find the

best motion vector which depicts the lowest distance of the current block to the reference block in terms of the sum of absolute error, mean square error or any other distance criteria.

14. (Original)   The video/audio processing method according to claim 13, wherein said motion estimation block performs motion estimation with variable block sizes.

15. (Original)   The video/audio processing method according to claim 13 including saving as a feature point list a feature point location, a block signature, a motion vector and the block distance for the best block position in a reference frame.

16. (Original)   The video/audio processing method according to claim 9, wherein said motion estimation block and said feature point tracking step include:

a step of performing motion estimation or feature point tracking in a discrete cosine transform domain for all search locations in the reference frame around the predicted motion vector in order to find the best motion vector which depicts the lowest distance of the current block to the reference block in terms of sum of absolute errors, mean square errors or any other distance criteria; and

a step of calculating the block signature in the DCT domain of the block having said best motion vector position.

17. (Original)   The video/audio processing method according to claim 16 including saving the feature point location, the block signature, motion vector and the block distance for the best block position in a reference frame as a feature point list.

18. (Original)   The video/audio processing method according to claim 1, wherein the motion vector and the block signature for all relevant current blocks are determined.

19. (Original)   The video/audio processing method according to claim 4, wherein said block relevance metric calculating step calculates a block relevance metric in the case when the current macro-block is an intra-type macroblock and the reference macroblock is a prediction coded macroblock or a bidirectionally prediction coded macroblock, said block relevance metric being calculated using a relevance measure as found based on the motion vector and the prediction error energy for an associated block by taking into account the reference macroblock.

20. (Original)   The video/audio processing method according to claim 4 including setting the block relevance metric to zero in the case when the current macroblock is a prediction coded macroblock or a bidirectionally prediction coded macroblock; and updating the list of already tracked feature points from the reference frame.

21. (Original)   The video/audio processing method according to claim 4 including calculating a block relevance metric in the case when the current macro-block is an intra-coded macroblock and the reference macro-block is also an intra-coded macroblock, said block relevance metric being calculated using a relevance measure as found based on the DCT activity from a block in the current macroblock and on the DCT activity as found by taking into account the reference macroblock.

22. (Original)   The video/audio processing method according to claim 1, wherein the video/audio signals are compression-encoded in accordance with MPEG1, MPEG2, MPEG4, DV, MJPEG, ITU-T recommendations H.261 or H.263.

23. (Original)   The video/audio processing method according to claim 1, wherein the extracted feature points are used along with metadata associated with these feature points for object motion estimation.

24. The video/audio processing method according to claim 1, wherein the extracted feature points are used along with metadata associated with these feature points for estimating the camera motion.

25. (Original)   The video/audio processing method according to claim 1, wherein the extracted feature points are used along with metadata associated with these feature points for calculating a motion activity model for video.

26. (Original)   The video/audio processing method according to claim 21, wherein said estimated camera motion is used to facilitate a transcoding process between one compressed video representation into an other compressed video representation.

27. (Currently Amended)   A video/audio signal processing apparatus for processing supplied compression-encoded video/audio signals, comprising:

means for parsing said video/audio signals <u>in a compressed domain of the video/audio signals</u> to extract therefrom motion vectors of said video/audio signals, DCT-coefficients and macroblock-type;

extraction means for using said extracted motion vectors, DCT-coefficients and macroblock-type to extract at least one compressed domain feature point representing characteristics of said video/audio signals in a compressed domain of said video/audio signals;

means for performing motion estimation of the extracted feature points; and

means for tracking the feature points associated with a motion vector through a pre-set number of frames of said video/audio signals,

wherein said extraction means calculates the block relevance metric of all blocks according to said DCT-coefficients in the current frame to determine a block having high relevance as a candidate of the feature point selected as the next feature point based on said motion estimation step.

28. (Canceled)

29. (Previously Presented)   The video/audio processing apparatus according to claim 27, wherein said extraction means includes means for performing inverse transform of transforming said compressed domain only for the blocks of high relevance selected by said metric calculating means and for performing motion compensation for a prediction coded macroblock or a bidirectionally prediction coded macroblock.

30. (Original) The video/audio processing apparatus according to claim 29, wherein said inverse transform is inverse discrete cosine transform.

31. (Original) The video/audio processing apparatus according to claim 28, wherein said current frame includes an arbitrarily shaped video object plane.

32. (Original) The video/audio processing apparatus according to claim 30, wherein said extraction means calculates and extracts a block signature for the current block of high relevance as selected in a discrete cosine transform domain using part or all of discrete cosine transform coefficients in a block.

33. (Original) The video/audio processing apparatus according to claim 30, wherein said extraction means calculates and extracts a block signature for the current block of high relevance as selected in a discrete cosine transform domain using part or all of individually weighted discrete cosine transform coefficients in a block.

34. (Original) The video/audio processing apparatus according to claim 30, wherein said extraction means calculates a block signature for the current block of high relevance as selected in a pel domain.

35. (Original) The video/audio processing apparatus according to claim 27, wherein said motion estimation means calculates an estimated motion vector, the position of a reference block and a search area in a reference frame.

36. (Original)   The video/audio processing apparatus according to claim 35, wherein said motion estimation means applies inverse transform of transforming said compressed domain to all blocks in an intra-macroblock in a search area of a reference frame.

37. (Original)   The video/audio processing apparatus according to claim 36, wherein said inverse transform is inverse discrete cosine transform.

38. (Original)   The video/audio processing apparatus according to claim 37, wherein said motion estimation means performs IDCT and motion compensation on all blocks in a prediction coded macroblock or in a bidirectional prediction coded macroblock in a search area of a reference frame.

39. (Original)   The video/audio processing apparatus according to claim 35, wherein said motion estimation means and said feature point tracking means performs motion prediction or feature point tracking in a pel area for all search locations in the reference frame around the predicted motion vector in order to find the best motion vector which depicts the lowest distance of the current block to the reference block in terms of the sum of absolute error, mean square error or any other distance criteria.

40. (Original)   The video/audio processing apparatus according to claim 39, wherein said motion estimation block performs motion estimation with variable block sizes.

41. (Original) The video/audio processing apparatus according to claim 39, wherein said motion estimation means and said feature point tracking means saves a feature point location, a block signature, a motion vector and the block distance for the best block position in a reference frame as a feature point list.

42. (Original) The video/audio processing apparatus according to claim 35, wherein said motion estimation block and said feature point tracking means performs motion estimation or feature point tracking in a discrete cosine transform domain for all search locations in the reference frame around the predicted motion vector in order to find the best motion vector which depicts the lowest distance of the current block to the reference block in terms of sum of absolute errors, mean square errors or any other distance criteria to calculate the block signature in the DCT domain of the block having said best motion vector position.

43. (Original) The video/audio processing apparatus according to claim 42, wherein said motion estimation block and said feature point tracking means saves the feature point location, the block signature, motion vector and the block distance for the best block position in a reference frame as a feature point list.

44. (Original) The video/audio processing apparatus according to claim 27, wherein the motion vector and the block signature for all relevant current blocks are determined.

45. (Original) The video/audio processing apparatus according to claim 30, wherein said block relevance metric calculating means calculates a block relevance metric in the

case when the current macro-block is an intra-type macroblock and the reference macroblock is a

prediction coded macroblock or a bidirectionally prediction coded macroblock, said block

relevance metric being calculated using a relevance measure as found based on the motion vector

and the prediction error energy for an associated block by taking into account the reference

macroblock.

46. (Original)   The video/audio processing apparatus according to claim 30,

wherein said extraction means sets the block relevance metric to zero in the case when the

current macroblock is a prediction coded macroblock or a bidirectionally prediction coded

macroblock and updates the list of already tracked feature points from the reference frame.

47. (Original)   The video/audio processing apparatus according to claim 30,

wherein said extraction means calculates a block relevance metric in the case when the current

macro-block is an intra-coded macroblock and the reference macro-block is also an intra-coded

macroblock, said block relevance metric being calculated using a relevance measure as found

based on the DCT activity from a block in the current macroblock and on the DCT activity as

found by taking into account the reference macroblock.

48. (Original)   The video/audio processing apparatus according to claim 27,

wherein the video/audio signals are compression-encoded in accordance with MPEG1, MPEG2,

MPEG4, DV, MJPEG, ITU-T recommendations H.261 or H.263.

00394755

49. (Original)   The video/audio processing apparatus according to claim 27, wherein the extracted feature points are used along with metadata associated with these feature points for object motion estimation.

50. (Original)   The video/audio processing apparatus according to claim 27, wherein the extracted feature points are used along with metadata associated with these feature points for estimating the camera motion.

51. (Original)   The video/audio processing apparatus according to claim 27, wherein the extracted feature points are used along with metadata associated with these feature points for calculating a motion activity model for video.

52. (Original)   The video/audio processing apparatus according to claim 50, wherein said estimated camera motion is used to facilitate a transcoding process between one compressed video representation into an other compressed video representation.

53. (Withdrawn)   A method for processing and browsing supplied video/audio signals, comprising the steps of:

building hierarchically a camera motion transition graph, wherein the graph building step includes the step of providing a graph layout having at least one main camera motion transition graph and having a plurality of nodes representing other camera motion with the transition paths illustrated for a video sequence;

browsing through the camera motion transition graph by depicting keyframes of a camera

motion video sequence at the nodes; and

browsing through the camera motion transition graph by depicting a graph representation

of the camera motion at the nodes.

54.   (Withdrawn)   An apparatus for processing and browsing supplied

video/audio signals, comprising:

means for building hierarchically a camera motion transition graph, wherein the graph

building means includes the step of providing a graph layout having at least one main camera

motion transition graph and having a plurality of nodes representing other camera motion with

the transition paths illustrated for a video sequence;

means for browsing through the camera motion transition graph by depicting keyframes

of a camera motion video sequence at the nodes; and

means for browsing through the camera motion transition graph by depicting a graph

representation of the camera motion at the nodes.

55.   (Withdrawn)   A method for extracting a hierarchical decomposition of a

complex video selection for browsing, comprising the steps of:

identifying video;

collecting key frames from said video shots for representing each video segment;

classifying the collections of key frames according to camera motion or global motion

information; and

building a graphical representation of the video, the graphical representation being based

upon the results of said classifying step, temporal as well as camera motion information

associated with each part of a video shot, wherein said graphical representation building step

includes the step of representing each category of video shot by node.

56. (Withdrawn)   An apparatus for extracting a hierarchical decomposition

of a complex video selection for browsing, comprising:

means for identifying video;

means for collecting key frames from said video shots for representing each video

segment;

means for classifying the collections of key frames according to camera motion or global

motion information; and

means for building a graphical representation of the video, the graphical representation

being based upon the results of said classifying step, temporal as well as camera motion

information associated with each part of a video shot, wherein said graphical representation

building step includes means for representing each category of video shot by node.

57. (Withdrawn)   A video/audio signal processing method for processing

supplied video/audio signals, comprising the steps of:

extracting at least one compressed domain feature point representing characteristics of

said video/audio signals in a compressed domain of said video/audio signals;

58. (Withdrawn)   The video/audio processing method according to claim 57, wherein said extraction step includes a step of calculating the block relevance metric of all blocks in the current frame to determine a block having high relevance as a candidate of the feature point selected as the next feature point based on said motion estimation step.

59. (Withdrawn)   The video/audio processing method according to claim 58, wherein said extraction step includes a step of performing inverse transform of transforming said compressed domain only for the blocks of high relevance selected by said metric calculating step and of performing motion compensation for a prediction coded macroblock or a bidirectionally prediction coded macroblock.

60. (Withdrawn)   The video/audio processing method according to claim 59, wherein said inverse transform is inverse discrete cosine transform.

61. (Withdrawn)   The video/audio processing method according to claim 58, wherein said current frame includes an arbitrarily shaped video object plane.

62. (Withdrawn)   The video/audio processing method according to claim 60 including calculating and extracting a block signature for the current block of high relevance as selected in a discrete cosine transform domain using part or all of discrete cosine transform coefficients in a block.

63. (Withdrawn) The video/audio processing method according to claim 60 including calculating and extracting a block signature for the current block of high relevance as selected in a discrete cosine transform domain using part or all of individually weighted discrete cosine transform coefficients in a block.

64. (Withdrawn) The video/audio processing method according to claim 60 including calculating a block signature for the current block of high relevance as selected in a pel domain.

65. (Withdrawn) The video/audio processing method according to claim 60, wherein said block relevance metric calculating step calculates a block reference metric in the case when the current macro-block is an intra-type macroblock and the reference macroblock is a prediction coded macroblock or a bidirectionally prediction coded macroblock, said block relevance metric being calculated using a relevance measure as found based on the motion vector and the prediction error energy for an associated block by taking into account the reference macroblock.

66. (Withdrawn) The video/audio processing method according to claim 60 including

setting the block relevance metric to zero in the case when the current macroblock is a prediction coded macroblock or a bidirectionally prediction coded macroblock; and

updating the list of already tracked feature points from the reference frame.

67. (Withdrawn)   The video/audio processing method according to claim 60 including

calculating a block relevance metric in the case when the current macro-block is an intra-coded macroblock and the reference macro-block is also an intra-coded macroblock, said block relevance metric being calculated using a relevance measure as found based on the DCT activity from a block in the current macroblock and on the DCT activity as found by taking into account the reference macroblock.

68. (Withdrawn)   The video/audio processing method according to claim 57, wherein the video/audio signals are compression-encoded in accordance with MPEG1, MPEG2, MPEG4, DV, MJPEG, ITU-T recommendations H.261 or H.263.

69. (Withdrawn)   A video/audio signal processing apparatus for processing supplied video/audio signals, comprising:

means for extracting at least one compressed domain feature point representing characteristics of said video/audio signals in a compressed domain of said video/audio signals.

70. (Withdrawn)   The video/audio processing apparatus according to claim 69, wherein said extraction means includes means for calculating the block relevance metric of all blocks in the current frame to determine a block having high relevance as a candidate of the feature point selected as the next feature point based on said motion estimation step.

00394755

71. (Withdrawn)   The video/audio processing apparatus according to claim 70, wherein said extraction means includes means for performing inverse transform of transforming said compressed domain only for the block of high relevance selected by said metric calculating means and of performing motion compensation for a prediction coded macroblock or a bidirectionally prediction coded macroblock.

72. (Withdrawn)   The video/audio processing apparatus according to claim 71, wherein said inverse transform is inverse discrete cosine transform.

73. (Withdrawn)   The video/audio processing apparatus according to claim 70, wherein said current frame includes an arbitrarily shaped video object plane.

74. (Withdrawn)   The video/audio processing apparatus according to claim 72 including

means for calculating and extracting a block signature for the current block of high relevance as selected in a discrete cosine transform domain using part or all of discrete cosine transform coefficients in a block.

75. (Withdrawn)   The video/audio processing apparatus according to claim 72 including

means for calculating and extracting a block signature for the current block of high relevance as selected in a discrete cosine transform domain using part or all of individually weighted discrete cosine transform coefficients in a block.

76. (Withdrawn)   The video/audio processing apparatus according to claim 72 including

means for calculating a block signature for the current block of high relevance as selected in a pel domain.

77. (Withdrawn)   The video/audio processing apparatus according to claim 72, wherein said block relevance metric calculating means calculates a block relevance metric in the case when the current macro-block is an intra-type macroblock and the reference macroblock is a prediction coded macroblock or a bidirectionally prediction coded macroblock, said block relevance metric being calculated using a relevance measure as found based on the motion vector and the prediction error energy for an associated block by taking into account the reference macroblock.

78. (Withdrawn)   The video/audio processing apparatus according to claim 72 including

means for setting the block relevance metric to zero in the case when the current macroblock is a prediction coded macroblock or a bidirectionally prediction coded macroblock; and

updating the list of already tracked feature points from the reference frame.

79. (Withdrawn)   The video/audio processing apparatus according to claim 72 including

means for calculating a block relevance metric in the case when the current macro-block is an intra-coded macroblock and the reference macro-block is also an intra-coded macroblock, said block relevance metric being calculated using a relevance measure as found based on the DCT activity from a block in the current macroblock and on the DCT activity as found by taking into account the reference macroblock.

80. (Withdrawn)   The video/audio processing apparatus according to claim 69, wherein the video/audio signals are compression-encoded in accordance with MPEG1, MPEG2, MPEG4, DV, MJPEG, ITU-T recommendations H.261 or H.263.

81. (Withdrawn)   A video/audio signal processing method for processing supplied video/audio signals, comprising the steps of:

performing motion estimation of at least one feature point representing characteristics of said video/audio signals in a compressed domain of said video/audio signals.

82. (Withdrawn)   The video/audio processing method according to claim 81, wherein said motion estimation step includes a step of calculating an estimated motion vector, the position of a reference block and a search area in a reference frame.

83. (Withdrawn)   The video/audio processing method according to claim 82 including applying inverse transform of transforming said compressed domain to all blocks in an intra-macroblock in a search area of a reference frame.

84. (Withdrawn)   The video/audio processing method according to claim 83, wherein said inverse transform is inverse discrete cosine transform.

85. (Withdrawn)   The video/audio processing method according to claim 84 including performing inverse IDCT and motion compensation on all blocks in a prediction coded macroblock or in a bidirectional prediction coded macroblock in a search area of a reference frame.

86. (Withdrawn)   The video/audio processing method according to claim 81, wherein the video/audio signals are compression-encoded in accordance with MPEG1, MPEG2, MPEG4, DV, MJPEG, ITU-T recommendations H.261 or H.263.

87. (Withdrawn)   A video/audio signal processing apparatus for processing supplied video/audio signals, comprising:

means for performing motion estimation of at least one feature point representing characteristics of said video/audio signals in a compressed domain of said video/audio signals.

88. (Withdrawn)   The video/audio processing apparatus according to claim 87, wherein said motion estimation means includes means for calculating an estimated motion vector, the position of a reference block and a search area in a reference frame.

89. (Withdrawn)   The video/audio processing apparatus according to claim 88 including

means for applying inverse transform of transforming said compressed domain to all blocks in an intra-macroblock in a search area of a reference frame.

90. (Withdrawn) The video/audio processing apparatus according to claim 89, wherein said inverse transform is inverse discrete cosine transform.

91. (Withdrawn) The video/audio processing apparatus according to claim 90 including

means for performing inverse IDCT and motion compensation on all blocks in a prediction coded macroblock or in a bidirectional prediction coded macroblock in a search area of a reference frame.

92. (Withdrawn) The video/audio processing apparatus according to claim 87, wherein the video/audio signals are compression-encoded in accordance with MPEG1, MPEG2, MPEG4, DV, MJPEG, ITU-T recommendations H.261 or H.263.